Review of Medical Image Retrieval Systems and Future Directions

Payel Ghosh, Sameer Antani, L. Rodney Long, George R. Thoma
U.S. National Library of Medicine, National Institutes of Health, Bethesda, MD-20852
{ghoshp3;santani;rlong;gthoma@mail.nih.gov}

Abstract

This paper presents a review of online systems for content-based medical image retrieval (CBIR). The objective of this review is to evaluate the capabilities and gaps in these systems and to determine ways of improving relevance of multi-modal (text and image) information retrieval in the iMedline system, being developed at the National Library of Medicine (NLM). Seven medical information retrieval systems: Figuresearch, BioText, GoldMiner, Yale Image Finder, Yottalook, Image Retrieval for Medical Applications (IRMA), and iMedline have been evaluated here using the system of gaps defined in [1]. Not all of these systems take advantage of the visual information contained in biomedical literature as figures and illustrations. However, all attempt to extract metadata about the image from the full-text of the articles and retrieve figures/images in response to a query. iMedline aims to advance the state-of-the-art in multimodal information retrieval by unifying image and text features in computing relevance. We discuss the shortcomings of these current systems and discuss future directions and next steps in iMedline toward context-based medical image retrieval.

1. Introduction

Biomedical information exists in different forms: as online literature databases such as PubMedCentral® [2] and BioMedCentral® [3], and as patients' cases in electronic health records (EHRs). This information, if effectively retrieved, could be useful for physicians, patients and those teaching and studying medical sciences for improving diagnosis, treatment planning, classroom learning and research. Online biomedical literature contains a rich source of visual information in the form of figures and illustrations that are not typically exploited by conventional bibliographic or full-text databases. Although captions and full-text excerpts contain descriptions of figures/illustrations, they cannot effectively represent the semantic information in medical images, which are better perceived visually by human experts. Our objective is to seek better ways to retrieve information from these sources by moving beyond conventional text-based searching and combining both text and visual features in search queries. As a first step towards this we review and evaluate seven medical information retrieval systems viz. FigureSearch, BioText, GoldMiner, Yale Image Finder, Yottalook, Image Retrieval for Medical Applications (IRMA), and NLM's iMedline. We then discuss the gaps and capabilities of these systems to determine ways of improving relevance of multi-modal (text and image) information retrieval in iMedline by including lessons learned from these efforts. Our broader goals are to improve retrieval of biomedical literature by targeting their visual content and retrieve semantically similar images to support differential diagnosis, clinical decision support, research and education.

This paper is organized as follows. We first provide an overview of current online medical information retrieval systems. This is followed by an evaluation of these systems using the semantic gaps defined in [1]: content, feature, usability, and performance gaps. We then describe future directions of research work for addressing these gaps and performing context-based medical image retrieval.

2. Medical Information Retrieval Systems

2.1 GoldMiner

The Goldminer® [4] searches figure captions to retrieve images from 11000 open-access peer-reviewed journal articles from the websites of American Roentgen Ray Society (ARRS), the American Society of Neuroradiology (ASN), the British Institute of Radiology (BIR), and the Radiological Society of North America (RSNA). It maps keywords in figure captions to concepts in NLM's Unified Medical Language System UMLS® metathesaurus and/or Medical Subject Heading (MeSH®) terms. Results are displayed in list or grid views as shown in Figure 1. Users have the option to search by age/modality/sex derived from the caption text. It also allows searches using multiple keywords.

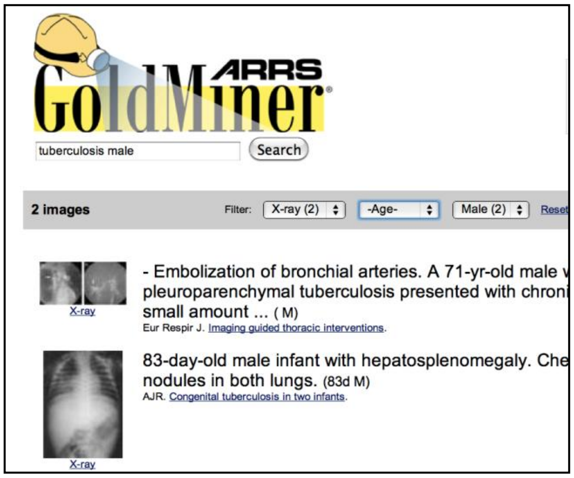

Figure 1. The ARRS GoldMiner search engine.

2.2 FigureSearch

FigureSearch search engine developed at the University of Wisconsin at Milwaukee, shown in Figure 2, is a component of the askHermes system [5], and is a tool devised to improve the quality of patient care by providing information to physicians at point of care. It uses the Lucene® text indexing and search technology to search online medical articles and generates a list view of results. Images are displayed on the left, while the title, authors, figure caption and summary are displayed on the right. The search engine separates itself from others with its ability to automatically generate summaries from papers (the purpose, experimental procedure, outcome and conclusion) using sentences from the main text.

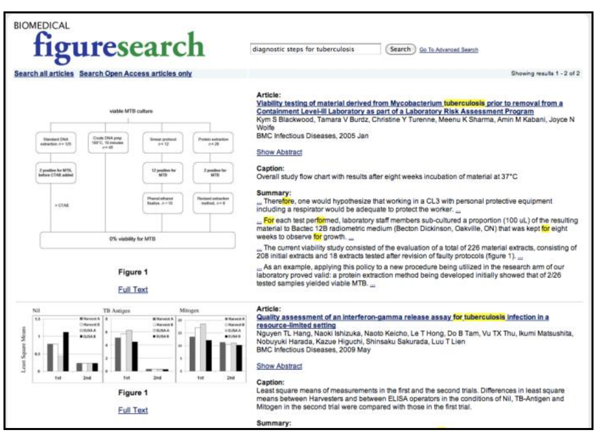

Figure 2. The FigureSearch search engine.

2.3 BioText

BioText search engine [6] developed at the University of California at Berkeley, shown in Figure 3, also uses the Lucene search engine to index over 300 open access journals and retrieves figures and text from online articles. Users can perform searches either on full-text or abstracts of journal articles. It is different from other search engines in that it can search table captions and retrieve part/expanded views of tables from online articles. Results can also be sorted by date and relevance.

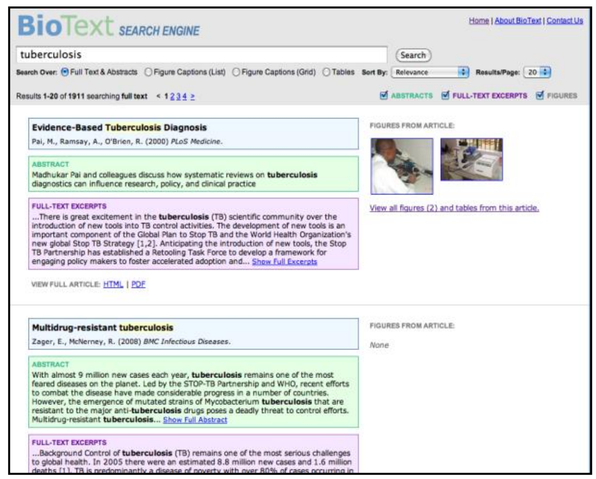

Figure 3. BioText search engine.

2.4 Yottalook

Yottalook [7] performs multilingual search in thirty three languages to retrieve images from peer-reviewed journal articles on the Web. It uses Google's indexing technology and a proprietary software called iVirtuoso for natural query analysis, semantic ontology generation and determining relevance. Natural query analysis generates keyword from search queries. Yottalook uses an enhanced version of the RSNA's RadLex® medical ontology to identify relationships or synonymous terms. This is known as semantic ontology generation. Relevance is automatically derived using a relevance algorithm (part of the iVirtuoso software) and is used to rank the retrieved results. Results can be viewed as grid or list views as shown in Figure 4 and allows users to save their searches using their myRSNA accounts.

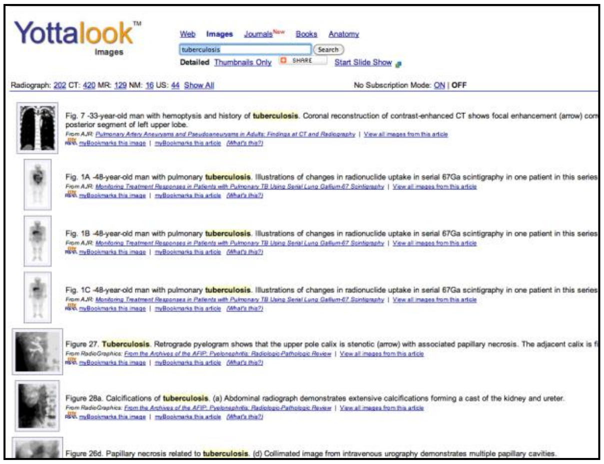

Figure 4. Yottalook search engine.

2.5 Yale Image Finder (YIF)

Yale Image Finder developed at Yale University, shown in Figure 5, searches text within biomedical images, captions, abstracts, and title to retrieve images from biomedical journal papers [8]. It uses optical character recognition to recognize text in images in both landscape and portrait modes and then validates the extracted text against content extracted from corresponding full-text articles. A unique capability of YIF is that users can access related images from associated papers by directly comparing image content. It uses Lucene® technology to index, search, and rank search results.

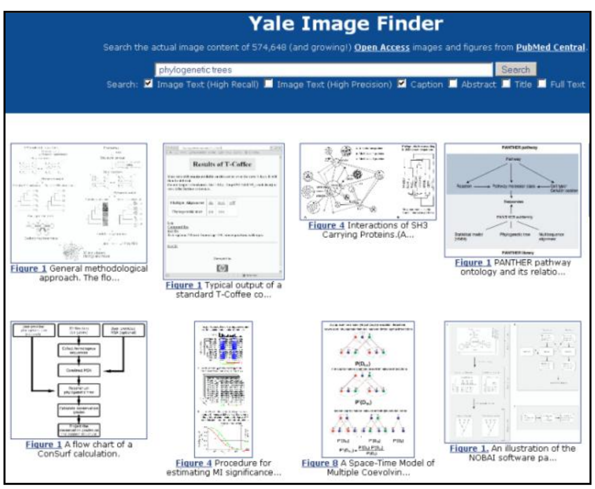

Figure 5. Yale image finder search engine.

2.6 IRMA

Image Retrieval for Medical Applications (IRMA) system, developed at Aachen University of Technology, Germany, aims to integrate text and image-based features for medical image retrieval [9]. The system, shown in Figure 6 indexes images using visual features and a limited number of text labels. Images are classified according to anatomy, bio-system, imaging direction, and modality of the image (x-ray, CT, MRI, etc.). It applies differential weighting of image features for computer-aided diagnosis. The image features are derived from co-registered training images. IRMA uses semantic layers to describe an image. These layers comprise multi-scale descriptions of the raw image data, extracted features, visual content and its spatial layout within the image. It supports text queries as well as image query by example (QBE) and has been tested on mammograms and bone x-ray images.

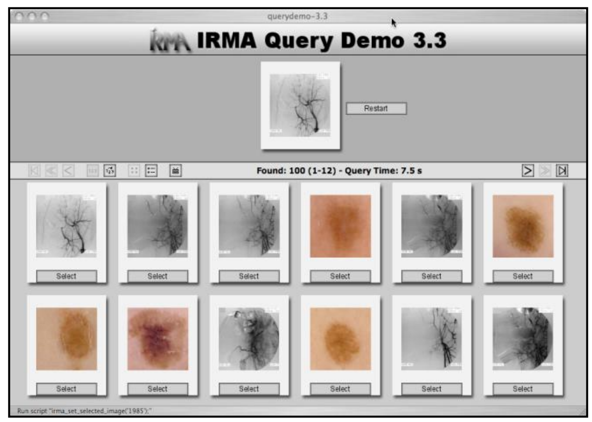

Figure 6. The IRMA image retrieval system.

2.7 iMedline

iMedline® is a multimodal search engine under development at NLM with goals to retrieve images from biomedical literature relevant to text and image queries and linking evidence automatically extracted from clinical articles to patients' cases. Along with the traditional elements of search results display, such as titles and author names, iMedline provides captions of the retrieved images and short summaries of the retrieved abstracts [10]. For the document retrieval task, iMedline uses NLM's Essie search engine. Essie is a phrase-based search engine with UMLS-based query expansion and probabilistic relevance ranking that exploits the document structure. The iMedline user interface provides the Essie search options and displays search results in grid or list views.

Figure 7 shows the snapshot of the iMedline search engine formerly known as Image and Text Search Engine (ITSE). In this system, retrieved images can be used as a query input to the relevance of search results. Currently, the image retrieval engine uses low-level visual features, such as color, texture, and shape as the primary building blocks of the visual content in an image. The features are then transformed into visual keywords, which annotate images with a set of labels that indicate the membership of local image regions/patches in various image categories. Similarity between a query image and database images is measured as a weighted linear combination of different features. The feature weights are updated to reflect the similarity rank of images judged by the user as relevant.

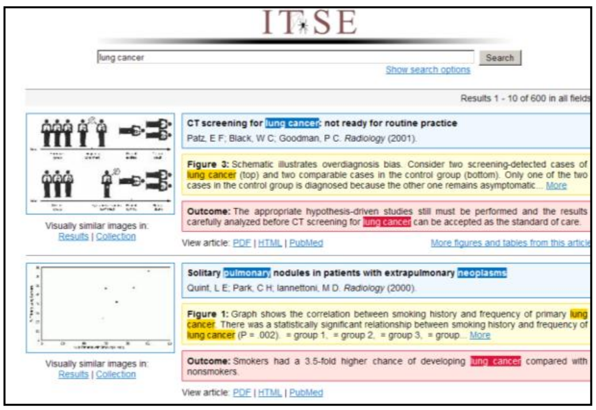

Figure 7. ITSE search engine displaying results in list mode.

2.8 Other Notable Systems

There are several other systems that use image features to perform medical image retrieval but have not been evaluated here because they are not online systems. For example, MedGIFT [11], VisMed [12], ASSERT [13], and BRISC [14]. MedGIFT was developed by the Viper group at the University of Geneva. It uses query-by-example to perform visual search. VisMed uses a visual vocabulary based on color and texture features to perform medical image retrieval. It was developed by the Institute for Infocomm Research in Singapore. ASSERT is a prototype medical image retrieval system developed at Purdue University, West Lafayette, IN. BRISC is a prototype CBIR system for lung nodule images developed by DePaul University, Chicago, IL, and University of Maryland, College Park, MD.

3. Evaluation of systems using gaps

Table 1 compares the CBIR systems described here using the criteria of gaps defined in [1]. Semantic gap refers to the difference in the level of image understanding at the human level versus the computer. Visual features derived from images contain low-level pixel information such as color, edge information, textures, etc. Mapping these features to high-level concepts such as spatial relationships between organs, identification of anatomical features, disease characterization is the main challenge faced by current CBIR systems. Four types of gaps were identified in CBIR systems in [1]: content, feature, usability, and performance gaps. Here we compare and contrast the current systems using this framework.

Table 1. Gaps addressed by current medical image retrieval systems as defined in [1].

		GoldMiner	Figure Search	BioText	Yottalook	YIF	IRMA	iMedline
Content	Semantic	UMLS concepts	Not addressed	Not addressed	Not addressed	Not addressed	Semantic layers of Blobworld	Visual keywords
	Context	Broad	Broad	Broad	Broad	Broad	Narrow	Broad
Feature	Type	Text-only	Text-only	Text-only	Text-only	Text-only	Text and image	Text and image
Performance	Use	Online	Online	Online	Online	Online	Online	Online
	Dataset	ARRS, ASN, BIR, and RSNA	PubMed, PubMed Central	Pubmed Central	750,000 images and figures	PubMed Central	10,000 radiographs	PubMed, PubMed Central
Usability	Query	Text	Text	Text	Text	Text	Text and Image	Text and Image
	Feedback	Not Addressed	Not Addressed	Not Addressed	Not Addressed	Not Addressed	Image	Image

3.1 Performance gap

Performance gap describes the scope of application of a search engine by comparing whether it can be used to retrieve information from a narrow/broad range of datasets. All the systems described here except the IRMA system have a broad operating scope and retrieve information from online journal databases. However, there is no method yet to perform a thorough quantitative evaluation of the retrieval performance of these systems on online biomedical journals.

3.2 Feature gap

Feature gap lists the type of text and/or image features used by a search engine. Currently text-only features are being used by GoldMiner, Figuresearch, YIF, Yottalook, and BioText. In contrast IRMA and iMedline use global image features such as color, and texture to represent the visual content of an image. However, gaps such as deriving multi-scale features, feature extraction from local regions-of-interest, region labeling, etc have not been addressed by these systems.

3.3 Usability gap

Usability gap is used to compare systems against types of queries and feedback used for improving relevance of search results. Current text-only information retrieval systems use keywords, phrases, and/or multiple keywords to perform retrieval. Systems using image features use both text and image queries. Gaps that need to be addressed in this domain are flexible query refinement, and improved relevance feedback. Query refinement techniques such as union, intersection and negation of queries along with hybrid queries that combine text and image features can be explored in future to improve the relevance of search results. User feedback and user interaction using relevance feedback also needs to be addressed by these systems.

3.4 Content gap

Semantic content gap refers to the text/image concepts used to perform information retrieval. The table clearly illustrates that systems using text-only information such as captions and full-text excerpts can only represent concepts that describe the content of an image to the extent of synonymous MeSH terms or UMLS concepts. They cannot represent the visual information present in images such as anatomical and pathology information, disease severity, etc., unless it is present in the text associated with an image. This drawback can be overcome by deriving visual information from medical images and mapping it to high-level textual medical concepts. The IRMA system demonstrates this capability on a narrow image domain of chest radiographs and mammograms. Our goal is to extend such a system and perform context-based image retrieval by combining text and image features. This involves research in the following areas:

(i) Pre-filtering to reduce the search space by automatic categorization of images using modality, body part, and orientation.

(ii) Identifying regions of interest in medical images using annotation markers within figures such as arrows, letters or symbols that are extracted from the image and correlating them with concepts in the associated text.

(iii) Generating effective ways to quantify the regions of interest/image patches so that they can be indexed and compared to perform similarity retrieval. In the next section, we describe our ongoing and future research efforts in these directions.

4. Future Work

4.1 Automatic Image Categorization

Automatic image categorization is a pre-filtering process, which can reduce the search space for faster and efficient similarity matching on large image collections. Medical images are categorized according to modality, body part, and orientation before visual similarity between images in the database and a query image are derived. Image features can then be extracted from the query image and compared against a set of indexed features to generate a ranked list of images based on a similarity score.

4.2 Automatic Image Annotation and ROI extraction

In order to extract meaningful information from images it is essential to derive features over relevant regions of interest or critical points within an image. Our approach to this problem is to first extract subfigures from composite figures and then finding useful "pointers" or annotations (arrows, symbols, or text labels) that point to the ROI [15]. We are also exploring methods to automatically detect and recognize annotations on images (arrows, text labels) as a means to correlate image ROIs with concepts extracted from the image captions [16].

4.3 Visual Keywords

In an effort to quantify images using visual features, we have developed "visual keywords" [17], or local image features and create a bag of concepts analogous to the bag of words representation commonly used in information retrieval from text documents. Visual keywords model the color and texture features derived from image "patches" generated by uniformly subdividing an image into non-overlapping regions. This approach has been used to automatically classify images into multiple modality categories. Visual keywords are also useful for improving the relevance of visually similar images using text-based image retrieval (IR) techniques. To extend the current framework to identify the gross anatomy of images, we are currently exploring unsupervised image segmentation techniques to characterize gross image regions [18].

To summarize, medical illustrations play a very important role in clinical decision making. Therefore, methods to effectively mine information from large databases of medical images using textual descriptions, image features and a combination of text and image features need to be explored. Most currently available search engines (Table 1) index the text associated with images such as captions and full-text excerpts to perform search in response to a user query. Our current efforts are geared towards performing semantic retrieval of biomedical images using a combination of text and image features applied to deriving regions of interest, representing regions or image patches as visual keywords and improving relevance feedback.

Acknowledgement

This research is supported by the Intramural Research Program of the National Institutes of Health (NIH), National Library of Medicine (NLM), and Lister Hill National Center for Biomedical Communications (LHNCBC).

5. References

[1] T. M. Deserno, S. Antani, R. Long, "Ontology of Gaps In Content-Based Image Retrieval", *Journal of Digital Imaging*, vol. 22, no. 2, pp. 202-15, April, 2009.

[2] PubMedCentral: http://www.ncbi.nlm.nih.gov/pmc/

[3] BioMedCentral: http://www.bmc.org

[4] ARRS GoldMiner: http://goldminer.arrs.org/

[5] Y. Hong, C. Yong-gang, "Automatically Extracting Information Needs from Ad Hoc Clinical Questions", *Proceedings of AMIA Symposium*, pp. 96-100, 2008.

[6] M. A. Hearst, A. Divoli, H. Guturu, A. Ksikes, P. Nakov, M. A. Wooldridge, J. Ye, "BioText Search Engine: beyond abstract search", *Bioinformatics*, vol. 23, no.16, pp. 2196-2197, 2007.

[7] Yottalook: http://www.yottalook.com/index_img.php

[8] S. Xu, J. McCusker, M. Krauthammer, "Yale Image Finder (YIF): a new search engine for retrieving biomedical images," *Bioinformatics*, vol. 24, no. 17, pp. 1968-1970, 2008.

[9] IRMA: http://ganymed.imib.rwth-aachen.de/irma/veroeffentlichungen_en.php

[10] iMedline: http://archive.nlm.nih.gov/iti/

[11] MedGIFT: http://medgift.hevs.ch/silverstripe/

[12] J. Lim, J. Chevallet, "VisMed: A Visual Vocabulary Approach for Medical Image Indexing and Retrieval", in proceedings of *Asia Information Retrieval Symposium*, pp. 84-96, 2005.

[13] ASSERT: http://cobweb.ecn.purdue.edu/RVL/Research/CBIR/index

[14] BRISC: BRISC: http://brisc.sourceforge.net/

[15] B. Cheng, J.R. Stanley, S. Antani, G.R. Thoma, "A Novel Computational Intelligence-based Approach for Medical Image Artifacts Detection", *Proceedings of the 2010 International Conference on Artificial Intelligence and Pattern Recognition*, pp.113-20, 2010.

[16] D. You, S. Antani, D. Demner-Fushman, M.M. Rahman, V. Govindaraju, G. R. Thoma, "Biomedical Article Retrieval Using Multimodal Features and Image Annotations In Region-based CBIR," *Proceedings of the Document Recognition and Retrieval Conference*. pp. 1-10, 2010.

[17] M. M. Rahman, S. Antani, G. R. Thoma, "A Classification-Driven Similarity Matching Framework For Retrieval of Biomedical Images", *11th ACM International Conference on Multimedia Information Retrieval*, pp.147-154, 2010.

[18] P. Ghosh, S. K. Antani, L. R. Long, G. R. Thoma, "Unsupervised Image Segmentation Using Cellular Automata for Region-based Image Retrieval", accepted for the *IEEE Conference on Healthcare Informatics, Imaging and Systems Biology*, 2011.